УДК 004.8

Ю.Г. Кривонос, Ю.В. Крак, О.В. Бармак, А.С. Тернов
Інститут кібернетики НАН України ім. В.М. Глушкова
Україна, 03187, м. Київ, просп. Академіка Глушкова, 40

Моделювання мімічної складової для українського жестового мовлення

Yu.G. Kryvonos, Iu.V. Krak, O.V. Barmak, A.S. Ternov
V.M.Glishkov Institute of Cubernetics of NASU
Ukraine, 03187, Kyiv, 40 Glushkova ave.

Modeling of Mimic Components for the Ukrainian Sign Language

Ю.Г. Кривонос, Ю.В. Крак, А.В. Бармак, А.С. Тернов
Институт кибернетики НАН Украины им. В.М. Глушкова
Украина, 03187, г. Киев, просп. Академика Глушкова, 40

Моделирование мимической составляющей для украинской жестовой речи

Розглянуто підхід для моделювання емоційної та артикуляційної складових обличчя віртуальної моделі людини в процесі відтворення нею українського жестового мовлення. Запропонований підхід дозволяє використовувати отриманні при моделюванні параметри для аналізу зображень обличчя людини в процесі відтворення жестового мовлення.
Ключові слова: віртуальна модель, жестове мовлення, ідентифікація емоцій та артикуляції.

An approach to mimics and articulation modeling on a virtual model of a human face in a process of reproduction of the Ukrainian sign language is proposed in this paper. The proposed approach allows using parameters, which were obtained during modeling, for analysis of images of a human face during sign language reproduction.
Key words: virtual model, sign language, emotion and articulation identification.

В статье предложен подход для моделирования эмоциональной и артикуляционной составляющей лица виртуальной модели человека в процессе воспроизведения украинской жестовой речи. Предложенный подход позволяет использовать полученные при моделировании параметры для анализа изображений лица человека в процессе воспроизведения жестовой речи.
Ключевые слова: виртуальная модель, жестовая речь, идентификация эмоций и артикуляции.

Вступ

Для синтаксису жестового мовлення характерна наявність немануального маркера: у запитальних реченнях, при загальному запитанні – підведені брови, при окремому (приватному) запитанні – опущенні брови та нахилена вперед голова. При запереченні – негативні рухи головою, відповідний вираз обличчя. Для розповідних речень характерна наявність на обличчі емоційного забарвлення, що відповідає смислу інформації, що передається. Крім того, у жестовому мовленні присутня (зазвичай) артикуляційна складова – промовляння за допомогою рухів губ відтворюваного жестами контенту звукової мови.

Отже, синтез мімічної складової (для моделювання процесу відтворення жестового мовлення за допомогою віртуальної моделі людини та за умови майбутнього аналізу (розпізнавання)) відіграє дуже важливу роль.

Метою даної роботи є розширення запропонованих технологій для моделювання та аналізу мімічних проявів емоцій [1] та аналізу міміки губ при промовлянні [2] для синтезу та аналізу мімічної складової для українського жестового мовлення. На основі технологій [1], [2] потрібно побудувати загальну модель, яка буде використовуватись як для відтворення жестового мовлення віртуальною моделлю людини, так і для розпізнавання відповідних маркерів на обличчі звичайної людини.

Міміка із засад концепції синтезу та аналізу жестового мовлення

Суттєві обмеження існуючих засобів відтворення жестової мови, з одного боку, та розвиток сучасної обчислювальної техніки, з іншого боку, спонукають до розробки більш гнучких алгоритмів, за допомогою яких можна було б створювати нові комп'ютерні системи навчання та комунікації для людей з вадами слуху. У роботах [3], [4] запропонована концепція інформаційної технології невербального спілкування людей з вадами слуху на базі використання мультимедійних можливостей сучасної комп'ютерної техніки для відтворення жестової мови.

Узагальнено концепція інформаційної технології зображена на рис. 1.

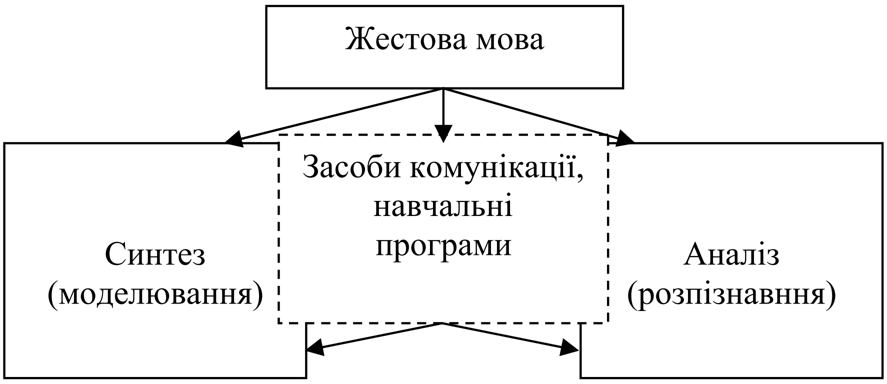


Рисунок 1 – Узагальнена концепція інформаційної технології

Основною ідеєю запропонованої концепції є синтез жестової мови через моделювання процесу її відтворення. З іншого боку, отримана при моделюванні жестової мови інформація використовується для розпізнавання (аналізу). Отримана таким чином інформація є вхідною (як окремо, так і в поєднанні) для створення нових засобів комунікації, навчальних програм тощо.

Синтезувати жестову мову пропонується за допомогою просторової моделі людини, яка дозволяє уникнути недоліків існуючих засобів відтворення жестової мови (фото- та відеозображення).

Просторова модель повинна синтезувати основні канали невербального спілкування: міміку, жести та пантоміміку. Виходячи з природи жесту як довільного знака, що продукується головою, руками, тілом та виражає емоції і повідомляє інформацію, потрібно запропонувати технології як для отримання та збереження жестів (послідовності зміни кутів скелета людини з плином часу), так і для формалізації мімічної складової (емоції тощо) (рис. 2).

Із запропонованої схеми (рис. 2), яка розкриває суть процесу синтезу жестової мови, з точки зору відображення на обличчі (як каналу невербальної інформації) випливають такі **основні задачі** (сигнали очей і брів можна звести до емоційної складової):

1) потрібно розробити інформаційну та математичну моделі для синтезу та аналізу емоційних проявів на обличчі;

2) потрібно розробити інформаційну та математичну моделі для синтезу та аналізу візуальних мімічних проявів при промовлянні усної мови (з урахуванням промовляння при виконанні жесту).

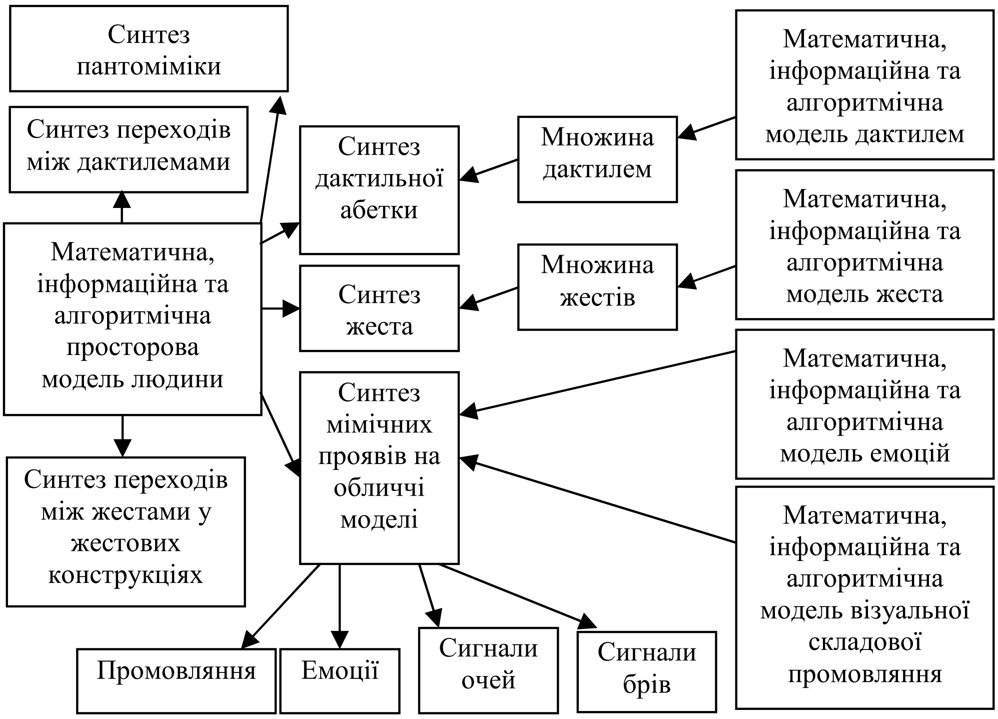


Рисунок 2 – Схема концепції просторової моделі для моделювання візуалізації жестової мови

Виходячи з того, що емоційні та артикуляційні прояви на обличчі людини мають одну природу – скорочення певних м’язів – пропонується розглядати задачу синтезу мімічних проявів комплексно: як для емоційної, так і для артикуляційної складових.

Формалізація складових, що впливають на формування мімічних проявів

Аналіз систематизації емоцій і їх виразів на обличчі, особливо методика FACS [5] дає змогу виділити ряд зовнішніх мімічних проявів на обличчі, суперпозиція яких формує емоційні та артикуляційні стани. Методика FACS оперує набором рухових одиниць обличчя, за допомогою яких відтворюються основні рухи, що здійснюються основними м’язами чи групами м’язів. Авторами методики FACS також розроблені емоційна система кодування рухів обличчя EmFACS [6]) і FACSAID [7], які розглядають тільки рухи обличчя, пов’язані з емоціями. За допомогою цих методик виведені відповідні формули (послідовності рухів (з певною інтенсивністю), суперпозиція яких відтворює певні емоції). За допомогою цих формул визначені такі емоції: здивування, страх, радість, смуток, відраза, зневага (презирство), гнів.

Після аналізу одиниць рухів м'язів обличчя для методики FACS були запропоновані ознаки, за допомогою яких можна синтезувати мімічні прояви емоцій, немануальні маркери синтаксису жестового мовлення та артикуляційні складові звукової мови. Основні з цих ознак (для брів та лоба) наведено у табл. 1.

Ці мімічні ознаки, подані у вигляді віднормованих значень, створюють систему координат для кодування довільної експресії чи артикуляції на обличчі. Початкова точка цієї системи координат відповідає стану спокою обличчя.

Таблиця 1 – Ознаки для синтезу мімічних проявів емоцій (для брів та лоба)

№	Ознака	Інтенсивність	
		Початкове значення	Кінцеве значення
1.	Брови (вертикаль)	Стан спокою	Підведені догори
2.	Брови (горизонталь)	Стан спокою	Зміщені до перенісся
3.	Куточки брів (зовнішні)	Стан спокою	Підведені догори
4.	Куточки брів (внутрішні)	Стан спокою	Підведені догори
5.	Лоб (вертикаль)	Стан спокою	Зморшки
6.	Лоб (горизонталь)	Стан спокою	Зморшки

Для формалізації емоцій пропонується перейти до вивчення ситуацій, в яких ці емоції виникають [1]. Тобто при визначенні емоцій в найбільш загальному вигляді описується ситуація, в якій вони виникають. Розрізнятимемо назву емоції і її позначення. Під позначенням будемо мати на увазі вектор (Em) (тобто абстрактне поняття) з такими ознаками:

$$Em_i^{\eta} = (\xi_1, \xi_2, \xi_3),\ i = \overline{1,8}, \tag{1}$$

де ξ – бінарні ознаки, які класифікують емоції: ξ_1 – ознака, яка визначає знак емоції – позитивна (1) емоція чи негативна (0); ξ_2 – ознака, яка визначає час виникнення емоції відносно події (передбачаюча (0) та констатуюча (1) емоції); ξ_3 – ознака, яка визначає направленість емоції. За цією ознакою виділяють емоції, направлені на себе (1) та направлені на зовнішні об'єкти, на інших людей (0).

Комбінуючи три описані бінарні ознаки, отримаємо 8 різних варіантів. Введемо 4-у ознаку (η). Ця ознака описує групи емоцій за джерелом їх походження:

η=1 –емоції, пов'язані з задоволенням (незадоволенням) особистих потреб людини;

η=2 – емоції, які виникають в результаті порівняння деякого об'єкта, самого себе або своїх дій зі своїми же нормами, стандартами, правилами, переконаннями;

η=3 – емоції, які виникають в результаті порівняння об'єкта з суспільними правилами та нормами;

η=4 – емоції, які виникають у зв'язку з потребами інших людей;

η=5 – емоції, які виникають у результаті взаємин з іншою людиною;

η=6 – емоції, які виникають на основі презирства.

Сполученням наведених ознак можна описати 48 якісно різних емоцій.

Метою подальшої формалізації є вибір назв емоцій, які найбільш точно задовольняють набору класифікуючих ознак. Після призначення кожній емоції чотирьох класифікуючих ознак цей набір стає її визначенням. Після формалізації емоцій у такий спосіб вони стають абстрактними об'єктами і з ними можна оперувати відповідно до їх означень, а не з особистим феноменологічним досвідом.

Далі, використовуючи запропоновані ознаки, означимо емоції для першої групи (емоції, які виникають на основі особистих потреб, η=1) [1]. **Горе** ($Em_3^1 = (0,1,1)$): під горем розумітимемо емоцію, яка виникає при втраті деяких цінностей. **Радість** **(**$Em_7^1 = (1,1,1)$): емоція, яка виникає в результаті задоволення деякої особистої потреби

у широкому сенсі. Типовою ситуацією яка приводить до радості буде ситуація досягнення бажаної мети. **Страх (** $Em_1^1 = (0{,}0{,}1)$): виникає в результаті передчуття втрати можливості задоволення деякої особистої потреби. **Надія (** $Em_5^1 = (1{,}0{,}1)$): виникає в результаті передчуття задоволення особистої потреби. **Гнів (** $Em_2^1 = (0{,}1{,}0)$): виникає в результаті незадоволення деякої особистої потреби, яка стимулює людину на подолання перешкоди, яка заважає її задоволенню. **Задоволення (** $Em_6^1 = (1{,}1{,}0)$): виникає в результаті досягнення деякої мети, пов'язаної з особистою потребою, направлене на припинення дій з досягнення цієї мети. **Інтерес (** $Em_4^1 = (1{,}0{,}0)$): емоція, яка виникає до об'єкта, з допомогою якого людина передбачає задовольнити свою потребу. **Зневага (** $Em_0^1 = (0{,}0{,}0)$): виникає в передчутті того, що об'єкт не задовольнить деяку нашу потребу. Якщо людина повела себе неправильно, то ми робимо висновок про її нездатність дати нам те, що потрібно, і можемо відчути до неї зневагу.

Розглянутий набір векторів-емоцій ($Em_i^1, i = \overline{0,7}$) будемо називати базовим, і будь-яку іншу емоцію будемо подавати у вигляді опуклої комбінації базових емоцій. Використовуючи означені вектори емоцій ($Em_i^1, i = \overline{0,7}$), побудуємо векторну модель емоцій за наступною схемою:

1) означимо емоції через різноманітні сполуки 4 ознак ситуацій, в яких вони виникають;

2) кожній емоції поставимо у відповідність деякий елемент векторного простору;

3) з допомогою визначень через ознаки ситуацій між векторами вводиться операція додавання;

4) множення на позитивне число моделює існування відносно більш сильних та більш слабких однакових емоцій;

5) множення на негативне число відображає факт існування протилежних емоцій.

В [1] доведено, що емоції можна представити у вигляді опуклої комбінації двох емоцій із вже розглянутих ($Em_i^1, i = \overline{0,7}$). Виходячи з цього маємо:

$$Em_i^\eta = \alpha Em_k^l + \beta Em_i^1, \alpha + \beta = 1, \beta > \alpha, \eta = \overline{2,6}, i = \overline{0,7}, k \in [0 \cdots 7], l \in [1 \cdots 6], \qquad (2)$$

де Em_i^η – i-а емоція для η-ї групи, Em_k^1 – емоція із 1-ї групи для формування зсуву при генерації емоцій для η-ї групи, Em_i^1 – емоція 1-ї групи, яка знаходиться на тій же вершині, що й емоція, яка генерується (тобто емоція, яка генерується, повинна мати такі ж ξ_1, ξ_2 і ξ_3, що й емоція з 1-ї групи, і вона повинна мати більшу вагу, ніж Em_k^1 (тому $\beta > \alpha$)).

Далі наведемо емоції, пов'язані з особистими нормами та правилами (η=2). Емоцією зсуву 1-ї групи є емоція задоволення (Em_6^1).

Провина ($Em_3^2 = \alpha Em_6^1 + \beta Em_3^1, \beta > \alpha$) є Задоволення (своїми принципами) + Горе (від необхідності нести відповідальність за порушення принципів. **Повага** ($Em_6^2 = \alpha Em_6^1 + \beta Em_6^1, \beta > \alpha$) є Задоволення + Задоволення (від того, що інша людина відповідає цим принципам). **Самоповага** ($Em_7^2 = \alpha Em_6^1 + \beta Em_7^1, \beta > \alpha$) є Задоволен ня + Радість (від відповідності самого себе цим принципам). **Презирство** ($Em_2^2 = \alpha Em_6^1 + \beta Em_2^1, \beta > \alpha$) є Задоволення + Гнів (направлений на подолання ситуації невідповідності цим принципам). **Симпатія** ($Em_4^2 = \alpha Em_6^1 + \beta Em_4^1, \beta > \alpha$) є Задоволення + Інтерес (до людини, яка, можливо, буде задовольняти потреби індивіда). **Антипа тія** ($Em_0^2 = \alpha Em_6^1 + \beta Em_0^1, \beta > \alpha$) є Задоволення + Зневага (до людини, яка, мож

ливо, не буде задовольняти потреби індивіда). **Почуття відповідальності** ($Em_5^2 = \alpha Em_6^1 + \beta Em_5^1, \beta > \alpha$) є Задоволення + Надія (на те, що індивід сам буде відповідати своїм нормам). **Безвідповідальність** ($Em_1^2 = \alpha Em_6^1 + \beta Em_1^1, \beta > \alpha$) є Задоволення + Страх (можливої невідповідності своїм нормам).

Аналогічним чином можна означити й емоції для інших η. Для емоцій, які виникають в результаті відповідності (невідповідності) чиїмось або суспільним стандартам, нормам, правилам *(η=3)*, емоцією зсуву 1-ї групи є емоція поваги (Em_6^2). Емоції, які виникають у зв'язку з чужими потребами *(η=4)*, мають емоцією зсуву 1-ї групи захоплення (Em_6^3). Емоціям, які виникають на основі взаємних відносин з іншими людьми *(η=5)*, зсувом 1-ї групи слугує емоція подяки (Em_6^4). Емоції на основі презирства *(η=6)* використовують, відповідно, для зсуву 1-ї групи емоцію презирства (Em_2^2).

Для переходу від моделі емоцій, що містить психологічні ознаки, до моделі з мімічними ознаками застосуємо підхід, запропонований для методики FACS [5-7]. Скористаємося запропонованим набором ознак для синтезу мімічних проявів емоцій (табл. 1). За допомогою цих ознак побудуємо векторний простір.

Екман [5] встановив, що існують сім основних виразів обличчя – конфігурацій (схем) міміки, які відтворюють сім емоцій: щастя, подив, страх, страждання, гнів, відраза (презирство) й інтерес. Екман виділив три автономні зони обличчя: область чола й брів, область очей (очі, повіки, основа носа) і нижня частина обличчя (ніс, щоки, рот, щелепи, підборіддя). Проведені дослідження дозволили розробити своєрідні «формули» мімічних виразів, які фіксують характерні зміни в кожній із трьох зон обличчя, а також сконструювати фотоеталони мімічних проявів ряду емоцій.

Щоб змоделювати емоційні вирази обличчя, потрібно спочатку визначити більш детальні їх залежності від рухів м'язів обличчя. У [5-7] наведена система для опису всіх візуально помітних рухів обличчя.

Використовуючи аналогічний підхід побудовані базиси векторного простору (з використанням набору ознак для синтезу мімічних проявів емоцій (табл. 1)). Для цілей синтезу та аналізу мімічних складових жестового мовлення запропоновані три базиси:

1) для основних емоційних станів (щастя, подив, страх, страждання, гнів, відраза й інтерес);

2) для немануальних маркерів синтаксису жестового мовлення (підняті брови, опущені брови тощо);

3) для артикуляційних мімік (15 візуальних портретів букв-візем та стан обличчя у спокої).

За допомогою запропонованих ознак можливо побудувати відповідні вирази обличчя (як лінійну комбінацію) для віртуальних моделей та, з використанням базисів, отримати розклад довільної міміки обличчя людини на відповідні значення ознак та на вклад кожного вектора із означених базисів.

Морфінг емоційних та артикуляційних складових для моделювання мімічних проявів на обличчі

У жестовому мовленні, окрім самого жесту, каналами передачі інформації є емоційна, немануальна синтаксична та артикуляційна складові. Для правильної інтерпретації відтворення жестового мовлення необхідно враховувати весь комплекс інформації, що супроводжує жест, бо один і той самий жест має багато значень в усному мовленні.

Для моделювання процесу анімації артикуляційної й емоційної складових використано механізм морфінгу – алгоритм плавного переходу від одного стану об’єкта до іншого. При застосуванні морфінгу використовуються тільки опорні стани, за допомогою яких розраховуються проміжні стани й моделюється процес анімації.

Відображення або побудова міміки на обличчі тривимірної моделі досягається при застосуванні відносного (сегментного) морфінгу до моделі. Нехай розглядаються рівнопотужні множини точок $X_0 \subset R^3$ та $X_i^{Morph} \subset R^3$, між точками яких існує строга відповідність. X_0 буде базовим мешем моделі або задаватиме базовий стан моделі, а X_i^{Morph} – морфом: модель зі зміненим базовим станом. Формула відносного морфінгу для *M* морфів у формалізмі моделі буде мати такий вигляд:

$$X = (w_0 X_0 + \sum_{m=1}^{M} w_m \cdot X_m^{Morph}) \Big/ \sum_{m=0}^{M} w_m \,, \qquad (3)$$

де w_m – вагові коефіцієнти. Якщо під моделлю розглядати модель голови людини, а під X_i^{Morph} розуміти зміни в базовому стані моделі при мімічних проявах, то результатом операції є лінійна комбінація мешів моделі й міміки.

Для кожної точки $x^k \in X$ і відповідних цій точці $x_i^k \in X_i^{Morph}$ при неврахуванні нерухомих точок формула (3) модифікується наступним чином:

$$x^k = (w_0 x_0^k + \sum_{m=1}^{M} w_m \cdot g_m \cdot x_m^k) \Big/ \sum_{m=0}^{M} w_m \cdot g_m \,,\ g_m = \begin{cases} 0, \left\| x_0^k - x_m^k \right\| < \partial \\ 1, \left\| x_0^k - x_m^k \right\| \geq \partial \end{cases}, \qquad (4)$$

де ∂ – поріг чутливості до зміни положення точки в просторі.

Нехай $g(a.b.x) = \begin{cases} 1, \ a < x \leq b \\ 0, \ інакше \end{cases}$, $a > b$. Тоді при використанні параметричного рівняння прямої $x = x^0(1-t) + x^1 t$ для моделювання процесу переходу з початкового стану у кінцевий загальне рівняння переходу зі стану X_1 до стану X_n через проміжні стани $X_2, X_3, ..., X_{n-1}$ в моменти часу $t_1, t_2, ..., t_n$, в яких модель має приймати відповідні положення X_i, матиме вигляд:

$$X = \sum_{i=1}^{n-1} g(t_i, t_{i+1}, t) \cdot (X_i \cdot \frac{1-(t-t_i)}{t_{i+1}-t_i} - X_{i+1} \cdot \frac{t_{i+1}-t}{t_{i+1}-t_i}) + X_1 \cdot g(t_1 - 1, t_1, t) \,, \qquad (5)$$

де $t \in [t_1, t_n]$.

Метод сегментного морфінгу дає можливість формувати кілька виразів обличчя на основі невеликої кількості морфів і змінювати стан обличчя при анімації мовлення. Додатковою перевагою цього морфінгу є те, що можна анімувати щелепу незалежно від губ та очей (моргання), незалежно від емоційного прояву на обличчі.

Для моделювання мімічних складових процесу анімації жестового мовлення важливо враховувати такі складові: артикуляційна віземa, емоція, немануальний маркер синтаксису жестового мовлення.

Артикуляційна візема – це характерний вираз обличчя, що є візуальним портретом фонеми або іншої базової звукової одиниці в розмовній мові. Використовуючи віземи, люди з порушеннями слуху сприймають розмовну мову візуально. Морфінг-

ціль віземи вважається одиницею візуальної інформації. Тому для розв’язання задачі синтезу правильної артикуляції при спілкуванні людей з дефектами слуху було проведено дослідження, спрямовані на отримання морфінг-цілей візем української мови [2].

В українській мові розрізняють 15 візем і стан спокою: 0) базовий стан моделі голови людини; 1) а; 2) е; 3) о; 4) у; 5) і, и; 6) й; 7) ш, ж, ч, дж; 8) к, г, х, ґ; 9) п, б, м; 10) в, ф; 11) т, д, н, л; 12) с, з, ц, дз; 13) р; 14) л', р'; 15) т', д', н'.

За тестовою вибіркою фонем української мови було отримано відеоряд процесу артикуляції професійного сурдоперекладача, на основі якого з урахуванням будови моделі голови людини були змодельовані морфінг-цілі артикуляційних візем української мови.

Для синтезу станів облич моделі з певними емоціями і немануальними маркерами створені морфінг-цілі, що відповідають наведеним вище ознакам. Відповідні лінійні комбінації цих морфінг-цілей відповідають конкретним емоціям.

Морфінг із використанням морфінг-цілей для фіксованого кадру здійснюється з урахуванням вагових коефіцієнтів:

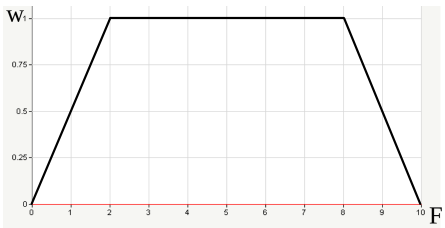
1) визначається кількість кадрів для анімації;

2) визначається набір емоцій, присутніх у слові (і/або немануальний маркер);

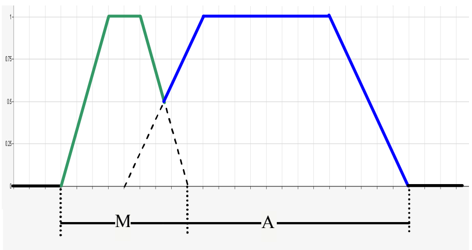
3) за фонетичною транскрипцією розраховується набір візем для візуалізації процесу артикуляції;

4) для візем розраховуються тривалості етапів анімації й точки появи віземи.

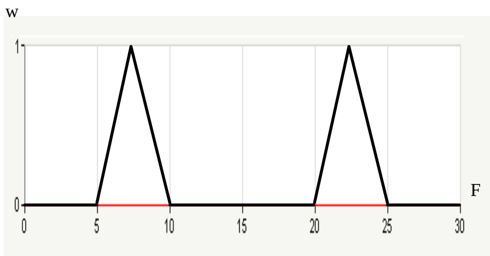
На основі отриманої інформації для кожного кадру будується набір артикуляційних візем й емоцій з їхніми ваговими коефіцієнтами. Формули розрахунку вагових коефіцієнтів подані графіками (рис. 3).



для однієї віземи (емоції)



для комбінації візем



для процесу моргання

Рисунок 3 – Алгоритм розрахунку вагових коефіцієнтів

Для відтворення процесу анімації артикуляційної й емоційної міміки тривимірною моделлю людини створено базовий клас, що реалізує морфемну анімацію. У класі реалізовані відповідні методи, які, використовуючи тривимірне API OPENGL, відтворюють за описаними параметрами модель людини й з використанням алгоритмів спінінгу (для жестової кінематики) й морфінгу (для міміки) реалізують анімацію. У кожному кадрі може бути присутня емоція, артикуляційна віземa або комбінація віземи і емоції.

Підходи до аналізу мімічних проявів на обличчі людини при жестовому мовленні

Для аналізу (розпізнавання) мімічних проявів на обличчі людини (за умови використання ознак запропонованого векторного простору для моделювання) потрібно виявити на зображенні значення ознак мімічних проявів (табл. 1). Для цього можливо застосувати різні підходи. Використовувались такі підходи (рис. 4): шаблони з NURBS-кривих [1-4], SURF-точки-особливості зображення [9], гнучкі шаблони для обличчя [8].



Шаблони з NURBS-кривих | SURF-точки-особливості | Гнучкий шаблон

Рисунок 4 – Застосування різних підходів для побудови вектора ознак мімічного стану обличчя

За допомогою зазначених підходів будувався базис B векторного простору, що складався з векторів, які відповідали базовим емоціям (радість, горе, надія, страх, задоволення, гнів, інтерес, зневага). Далі визначався вектор b, що відповідав конкретній емоції. Значення елементів вектора x

$$x = \left(B^T B\right)^{-1} \cdot B^T b \tag{6}$$

вказували на конкретний вклад кожної з базових емоцій у довільній емоції b, при цьому умова $\det\left(B^T B\right) \neq 0$ завжди досягалась.

Для аналізу артикуляційної мімічної складової застосовувався аналогічний підхід. Будувався базис з 16 артикуляційних складових (15 артикуляційних візем і стану спокою). Отриманий базис був недостатньо лінійно-незалежним – не виконувалась умова $\det(B^T B) > \varepsilon > 0$. Це випливає з того, що артикуляційні складові досить подібні. У такому випадку найбільш надійним методом для розв'язання задачі є метод сингулярного розкладу (*SVD)*. Було введено поріг τ близькості до нуля сингулярних чисел, який відображає помилки в початкових даних та обчисленнях. Тоді наближений розв'язок задачі (7) шукається наступним чином:

$$x = B^{+} b = V \sum{}' U^T b\,, \tag{9}$$

де $\sum{}' = diag\left(\sigma_1', \sigma_2', \cdots, \sigma_n'\right)$, $\sigma_j' = \dfrac{1}{\sigma_j}$, для $\sigma_j \geq \tau$ і $\sigma_j' = 0$ для $\sigma_j < \tau$; V , U^T – матриці з ортонормованими стовпцями. Результатом є вектор розкладу x, на основі якого приймається рішення про відповідність вхідного вектора конкретним базовим мімікам при промовлянні українською мовою.

Висновки

У даній роботі запропонований комплексний підхід для моделювання та розпізнавання мімічної складової для українського жестового мовлення. Синтез мімічної складової дозволяє досить реалістично відтворювати емоційну, артикуляційну та немануальну експресії на обличчі віртуальної моделі при відтворенні нею українського жестового мовлення. Отримана при моделюванні інформація (ознаки мімічних станів) використовується для розпізнавання відповідних мімічних складових на обличчі реальної людини при відтворенні нею українського жестового мовлення. Подальші дослідження будуть направлені на удосконалення методів запропонованого підходу.

Література

1. Кривонос Ю.Г. Моделювання та аналіз мімічних проявів емоцій / Ю.Г. Кривонос, Ю.В. Крак, О.В. Бармак [та ін.] // Доповіді НАНУ. – 2008. – № 12. – С. 51-55.
2. Кривонос Ю.Г. Розпізнавання міміки губ при промовлянні слів українською мовою / Ю.Г. Кривонос, Ю.В. Крак, О.В. Бармак [та ін.] // Доповіді НАНУ. – 2010. – № 5. – С. 41-44.
3. Кривонос Ю.Г. Інформаційна технологія для моделювання української мови жестів / Ю.Г. Кривонос, Ю.В. Крак, О.В. Бармак [та ін.] // Штучний інтелект. – 2009. – № 3. – С. 186-197.
4. Кривонос Ю.Г. Інформаційна технологія невербального спілкування людей з вадами слуху / Ю.Г. Кривонос, Ю.В. Крак, О.В. Бармак [та ін.] // Штучний інтелект. – 2008. – № 3. – С. 325-331.
5. P. Ekman. Facial Action Coding System: A Technique for the Measurement of Facial Movement / P. Ekman, W. Friesen. – Consulting Psychologists Press, Palo Alto, 1978.
6. Friesen W. EMFACS-7: Emotional Facial Action Coding System. Unpublished manual / W. Frisen, P. Ekman. – California : University of California, 1983.
7. Facial Action Coding System Affect Interpretation Dictionary (FACSAID) [Електронний ресурс]. – Режим доступу : http://www.face-and-emotion.com/dataface/facsaid/description.jsp
8. Saragih J. Deformable Model Fitting by Regularized Landmark Mean-Shifts / J. Saragih, S. Lucey, J. Cohn // International Journal of Computer Vision (IJCV). – 2010.
9. Bay H. SURF: Speeded Up Robust Features / Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool // Computer Vision and Image Understanding (CVIU). – 2008. – Vol. 110, № 3. – P. 346-359.

Literatura

1. Krivonos Y.G. Dopovidi NANU. 2008. № 12. S. 51-55.
2. Krivonos Y.G. Dopovidi NANU. 2010. № 5. S. 41-44.
3. Krivonos Y.G. Iskusstvennyj intellect. 2009. № 3. S. 186-197.
4. Krivonos Y.G. Iskusstvennyj intellect. 2008. № 3. S. 325-331
5. P. Ekman and W. Friesen. Facial Action Coding System: A Technique for the Measurement of Facial Movement. Consulting Psychologists Press, Palo Alto, 1978.
6. Friesen, W.; Ekman, P. (1983). EMFACS-7: Emotional Facial Action Coding System. Unpublished manual, University of California, California.
7. Facial Action Coding System Affect Interpretation Dictionary (FACSAID) http://www.face-and-emotion.com/dataface/facsaid/description.jsp
8. J. Saragih, S. Lucey and J. Cohn, Deformable Model Fitting by Regularized Landmark Mean-Shifts, International Journal of Computer Vision (IJCV), 2010.
9. Bay H. SURF: Speeded Up Robust Features" / Herbert Bay , Andreas Ess, Tinne Tuytelaars, Luc Van Gool, // Computer Vision and Image Understanding (CVIU). - 2008. Vol. 110, No. 3, p. 346-359.

RESUME

Yu.G. Kryvonos, Iu.V. Krak, O.V. Barmak, A.S. Ternov

Modeling of Mimic Components for the Ukrainian Sign Language

Modeling of a mimic component for the Ukrainian sign language is described in this paper. Mimic component plays an important or even crucial role in sign language of impaired hearing people. It should be noted that non-manual marker is a characteristic part of sign language syntax. In interrogative sentences in general question, it is raised eyebrows; in particular question, it is lowered eyebrows and head tilted forward. In negation, the negative movements of a head and corresponding facial expression take place. Emotions shown on a human face that corresponds to the information is a usual part of declarative sentences. In addition, there is usually an articulation component for pronunciation with a help of lips movement of a voice content of a gesture which is reproduced.

The authors propose a concept of information technology for non-verbal communication of people with impaired hearing. The basic idea of the proposed conception is synthesis of sign language through the modeling of its reproduction process. On the other hand, the obtained information during modeling of sign language is used for recognition (analysis). Information obtained in such way is input (both separately and in combination) for creation of new means of communication, training programs.

An approach to mimics and articulation modeling on a virtual model of a human face in the process of reproduction of the Ukrainian sign language is proposed in this paper. The proposed approach allows using parameters, which were obtained during modeling, for analysis of images of a human face during sign language reproduction.

Стаття поступила в редакцію 02.07.2012.